Contemporary Mathematics
https://ojs.wiserpub.com/index.php/CM/

Research Article

Exploring the Effectiveness of Regularized Representation-Based Feature Selection for PCR in Different Years with THz-TDS

Xiaowen Zha[1], Chengyong Zheng[1*], Zihao Liu[1], Qian Li[2], Zhijing Ye[3]

[1] School of Mathematics and Computational Science, Wuyi University, Jiangmen, 529000, China
[2] China Communication Technology(Jiangmen) CO., Ltd., Jiangmen, 529000, China
[3] Faculty of Science and Technology, University of Macau, Taipa, Macau
E-mail: zcy@wyu.edu.cn

Received: 15 Otober 2025; **Revised:** 13 November 2025; **Accepted:** 21 November 2025

Abstract: Pericarpium Citri Reticulatae (PCR) of longer storage years generally possesses higher medicinal value. However, it often challenging to distinguish them through traditional sensory analysis methods. In recent years, Terahertz Time-Domain Spectroscopy (THz-TDS) technology has been widely applied in food and drug identification, and feature selection methods based on regularized representation have also been extensively investigated. Therefore, this paper probes the feasibility of feature selection methods based on regularized representation for identifying PCR in different years by using THz-TDS. The average accuracy and standard deviation are adopted to assess the performance of the feature selection methods. The main exploration process is as follows: Firstly, we make use of the absorption coefficient and refractive index spectra with the frequency ranges of 0.1-1.0 THz and 0.1-1.5 THz respectively. Secondly, the Support Vector Machine (SVM) is employed as the classifier to establish classification methods in combination with feature selection methods based on regularized representations, which are then applied to identify PCR in different years. Furthermore, we propose a feature selection method based on sparse Low-Rank and Graph Structure Learning (LRGSL), which combines the advantages of feature selection methods based on regularized representation in recent years. The experimental results indicate that the feature selection method based on regularization representation is conducive to enhancing the accuracy of identifying the storage years of PCR, further demonstrating the feasibility of THz-TDS technology combined with the feature selection method based on regularization representation. Moreover, LRGSL, Joint Low-Rank Decomposition and Local Preservation (JLRDLP), and Low-Rank approximation and Structural Learning (LRSL) methods exhibit superior performance in the discrimination of PCR storage years.

***Keywords*:** Terahertz Time-Domain Spectroscopy (THz-TDS), Pericarpium Citri Reticulatae (PCR), regularized representation, feature selection, classification

MSC: 65K10, 68T05, 62H30

1. Introduction

Pericarpium Citri Reticulatae (PCR) is derived from the dried mature Pericarp of Citrus reticulata Blanco (Rutaceae) [1]. PCR is a famous herbal medicine and functional food and is officially listed in the Chinese Pharmacopoeia. Because

Copyright ©2026 Chengyong Zheng, et al.
DOI: https://doi.org/10.37256/cm.7220268914
This is an open-access article distributed under a CC BY license
(Creative Commons Attribution 4.0 International License)
https://creativecommons.org/licenses/by/4.0/

its main active ingredients are flavonoids, volatile oils and alkaloids [2], it is widely used clinically as folk medicine to treat indigestion, inflammatory respiratory diseases and cough [3]. The significance of bioactive compounds in citrus peels is further underscored by recent research, such as the work by Husni et al. [4], which demonstrated the antioxidant and wound-healing activities of essential oils from other Citrus species, reinforcing the broader medicinal value of this plant family. In addition, PCR is also widely used as a food ingredient in dietary condiments, tea and cakes, due to its unique aroma and health benefits [5].

In recent years, studies have shown that with the extension of storage years, the content of flavonoid components increases significantly [6]. During the aging process, the bound phenols and flavonoids undergo significant increases, accompanied by an increase in their antioxidant activity, as measured by various antioxidant analysis methods [7]. Therefore, the medicinal value of PCR is highly correlated with the storage year. However, PCR with different storage years have similar morphological characteristics, making it difficult for consumers to identify and distinguish the storage year visually, which often leads to the phenomenon of falsely labeled storage years of PCR in the market.

The conventional method for identifying PCR in different years is based on sensory analysis, mainly by examining its differences in appearance, taste, and aroma. However, this detection method is highly subjective. Currently, there are also some instrumental analysis techniques, including High-Performance Liquid Chromatography (HPLC) and Gas Chromatography-Mass Spectrometry (GC-MS) [8], that can be used to analyze the active components in PCR to determine their age. However, these analysis techniques are usually complex, time-consuming, and require chemical reagents. Therefore, developing new technologies to effectively and efficiently identify and distinguish PCR in different years has significant value.

Terahertz (THz) spectroscopy has emerged as a cutting-edge non-destructive identification technique in recent years. Classified as electromagnetic waves spanning the frequency range of 0.1-10 THz, THz radiation occupies the spectral region between microwaves and infrared rays [9]. Among the widely used THz spectroscopic approaches, THz Time-Domain Spectroscopy (THz-TDS) stands out for its ability to not only capture both amplitude and phase data of THz signals interacting with samples but also calculate key spectral parameters-specifically the absorption coefficient and refractive index-of the tested materials [10]. Distinct from alternative spectroscopic methods, THz spectroscopy exhibits unique detection attributes, including a broad spectral bandwidth, strong penetrability, high coherence, low photon energy, and material-specific "fingerprint" spectral responses [11]. Owing to these advantages, it has been recognized as a highly promising emerging detection tool, with applications ranging from foreign object detection and edible oil differentiation to the quantification of pesticide and antibiotic residues in agricultural commodities and the identification of genetically modified food products [12].

The spectral data collected from THz-TDS for samples contain rich physical and chemical information. However, due to the interference of background and random noise, many useful feature information is often concealed. Feature selection, as one of the important methods of data processing, aims to screen out features closely related to the classification target from the original data, while eliminating irrelevant or redundant features. This process serves to reduce data dimensionality, simplify the models, and improve the recognition accuracy and stability of the models [13]. In the field of machine learning, regularization is a powerful technique for improving the generalization ability of models and reducing the risk of overfitting. Especially when the number of features is large or there exists multicollinearity among features, the feature selection method with regularization becomes particularly important. By incorporating penalty terms into the loss function, regularization can not only control the complexity of the model but also promote the sparsity of model coefficients, thus achieving the purpose of feature selection. The advantage of this method lies in its ability to identify and select the most critical variables for the prediction target from numerous features, while simplifying the model structure and enhancing the interpretability and efficiency of the model.

In recent years, feature selection methods based on regularization representation have received widespread attention. Yang et al. [14] proposed an Unsupervised Discriminative Feature Selection (UDFS) algorithm based on $l_{2,1}$-norm regularization, which combines the $l_{2,1}$-norm and discriminant analysis to select discriminative feature subsets. The UDFS algorithm demonstrates that the $l_{2,1}$-norm achieves excellent results in structural sparsity. Zhu et al. [15] presented an unsupervised feature selection algorithm based on Regularized Self-Representation (RSR), which leverages the $l_{2,1}$-norm to describe the feature coefficient matrix and employs the residual matrix to eliminate redundant data,

thereby enhancing the model's robustness. Although the RSR method effectively utilizes the sparsity of the $l_{2,1}$-norm, it lacks exploration of the true correlation between data features. To address this issue, Wang and Wang [16] developed an unsupervised feature selection method founded on Low-Rank approximation and Structural Learning (LRSL)-this approach enhances model performance by integrating structural learning and sparse regularization into the baseline feature selection framework. Although the LRSL method incorporates low-rank regularization, the parameter k (which defines the rank) is challenging to predefine. To approximate the true low-rank property of the samples, Wang et al. [17] introduced a new unsupervised feature selection algorithm based on Low-Rank Sparse Self-Representation (LRSSR), utilizing the F-norm to measure the residual matrix of feature self-representation and the $l_{2,1}$-norm as a sparse regularization term for the feature coefficient matrix. Additionally, to better discover the true low-rank structure of the data, the Schatten P-norm is added to the feature coefficient matrix as a low-rank constraint. Experimental results in [17] show that the Schatten P-norm can better exploit the low-rank structure between data. Later, the Schatten P-norm was also introduced into supervised feature selection methods by Lim et al. [18], leading to the proposed supervised feature selection algorithm (Low-Rank Feature Selection: LRFS) based on sparse low-rank dual constraints. The LRFS method explores the correlation between features by considering the sparsity and low-rankness of the weight matrix in the objective function. Jiang and Chen [19] proposed a spectral feature selection algorithm (Joint Low-Rank Decomposition and Local Preservation: JLRDLP) based on low-rank decomposition and local preservation, which predicts pseudo-labels F using spectral clustering. The JLRDLP method reveals the global correlation of sample features by decomposing the feature coefficient matrix into the product of two low-rank matrices and adopting a graph structure in the objective function, and considers the sparsity of the feature coefficient matrix using the $l_{2,1}$-norm.

Despite the numerous feature selection methods based on regularized representation, there have been no reported cases of their application to food and drug identification using THz-TDS. Therefore, this paper is the first to apply feature selection methods for regularized representation to feature selection of THz spectral data, and explore their effectiveness in the discrimination of PCR in different years. It also provides an objective comparison of feature selection methods. In addition, this paper proposes a novel supervised feature selection model based on Low-Rank and Graph Structure Learning (LRGSL), combining the sparsity of the $l_{2,1}$-norm, the low-rank property of the Schatten P-norm, and the Laplacian graph structure. An iterative solution algorithm based on the Alternating Direction Method of Multipliers (ADMM) is also introduced. LRGSL fully considers the sparsity and low-rankness of the feature sparse matrix in the objective function, while also conducting sparse and structural learning of the correlations between features.

The main contributions are the following:

(1) Addressing the challenge of identifying the storage year of PCR, this paper explores the feasibility of using THz-TDS combined with feature selection methods based on regularized representation for identifying PCR in different years.

(2) The exploration process is elaborated, including selecting absorption coefficient and refractive index spectra from different frequency ranges, and using SVM as a classifier to establish classification methods with different feature selection methods based on regularized representation.

(3) A feature selection method based on LRGSL is proposed, which combines the advantages of recent regularized representation feature selection methods.

(4) Experiments were conducted to demonstrate that the regularized representation-based feature selection method effectively improves the accuracy of PCR discrimination using THz-TDS technology. A comparative analysis of the precision achieved by different feature selection methods was also presented.

The following content is organized as follows: Section 2 introduces related works, Section 3 presents the LRGSL algorithm, Section 4 describes the experimental preparation and methodology, Section 5 provides the experimental results and analysis, and the final section concludes on the paper.

2. Related works

To maintain consistency in notation throughout this paper, this section first defines some symbols and then introduces several feature selection methods based on regularized representations, including RSR [15], LRSL [16], LRSSR [17], LRFS [18], and JLRDLP [19].

2.1 *Symbols and definitions*

Let $X \in R^{n\times d}$ be the THz spectral data to be processed, where n is the number of samples, d is the dimension of each sample. Let $x_i = [x_{i1},\ x_{i2},\ \cdots,\ x_{id}] \in R^{1\times d}$ $(i = 1,\ 2,\ \cdots,\ n)$ be the i-th sample, i.e., the i-th row of X, $x^j = [x_{1j},\ x_{2j},\ \cdots,\ x_{nj}]^T \in R^{n\times 1}$ $(j = 1,\ 2,\ \cdots,\ d)$ be the data of the j-th feature, i.e., the j-th column of X, $y_i \in \mathscr{Y}=\{1,\ \cdots,\ c\}$ $(c \geq 2)$ be the label of x_i, c is the total number of classes. Let $Y \in R^{n\times c}$ be the sample label matrix after applying one-hold encoding. Assume the feature domain of X is $\Omega = \{1,\ 2,\ \cdots,\ d\}$. The feature selection method aims to find an index subset $\Omega_{opt} \subset \Omega$ so that the data with the selected features, denoted by $\tilde{X}=(x_{\cdot,\ j})_{|j\in\Omega_{opt}}$, has the best classification performance.

The Frobenius norm (F-norm) (denoted as $\|\cdot\|_F$) of matrix X is defined as

$$\|X\|_F = \sqrt{\sum_{i,\ j}|x_{ij}|^2} = \sqrt{\sum_{i=1}^{n}\|x_i\|_2^2} = \sqrt{tr(X^TX)}. \tag{1}$$

The F-norm measures the total energy or magnitude of the entire spectral data matrix X. In the context of THz spectroscopy, it can be interpreted as a cumulative measure of the absorption or reflection intensity across all samples and all frequency points (features). When used in an optimization objective (e.g., as a loss function in Eq. (7) and Eq. (8)), minimizing the F-norm of a residual matrix (like $X - XW$ or $Y - XW$) implies that we are seeking a model that reconstructs the original data or labels with the smallest overall squared error. This provides a globally accurate approximation.

The nuclear norm of X is defined as

$$\|X\|_* = tr(\sqrt{X^TX}), \tag{2}$$

where $tr(X^TX)$ represents the trace of matrix X^TX. The nuclear norm serves as a convex surrogate for the matrix rank. In signal processing, a low-rank matrix implies that the data points (rows) or features (columns) are highly correlated and lie on a lower-dimensional subspace. For our THz spectral data, imposing a low-rank constraint on the feature coefficient matrix W (as in Eqs. (7), (8), (9)) suggests that we believe there exists a latent low-dimensional structure governing the relationships between the original high-dimensional features. By minimizing $\|W\|_*$, we encourage W to capture these fundamental correlations, thereby filtering out noisy or redundant features and improving the model's generalization.

The $l_{2,\ 1}$-norm of X is defined as

$$\|X\|_{2,\ 1} = \sum_{i=1}^{n}\|x_i\|_2. \tag{3}$$

The $l_{2,\ 1}$-norm is a critical tool for feature selection. It is computed by first taking the l_2-norm of each row (i.e., the magnitude of each sample's vector), and then summing these magnitudes across all rows. When applied to a matrix like the feature coefficient matrix W (where each row corresponds to a feature's influence across all outputs), this norm induces row-sparsity. Conceptually, it penalizes the model based on the number of features used rather than the sheer magnitude of coefficients. A feature that is unimportant will have a row of coefficients with a very small l_2-norm, effectively pushing

the entire row to zero. Thus, after optimization, only the rows corresponding to the most discriminative features remain non-zero, achieving the goal of feature selection. In our THz data, this means identifying the specific frequency points that are most informative for classification.

2.2 *Algorithms to be compared*

RSR algorithm [15]: RSR is a classic unsupervised feature selection method. It employs a feature self-representation methodology to explore relationships among features. Through the utilization of $l_{2,\,1}$-norm to characterize the eigencoefficient matrix and residual matrix, this algorithm effectively eliminates redundant information and enhances the robustness of the model. The optimization problem of the RSR algorithm can be formulated as follows:

$$\min_{W} \|X - XW\|_{2,\,1} + \beta \|W\|_{2,\,1}, \tag{4}$$

where $W \in R^{d\times d}$ is the feature coefficient matrix, and β is the regularization parameter that controls the row sparsity of W. The first term ensures

$$X \approx \left(x^1,\, x^2,\, \ldots,\, x^d\right) \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_d \end{pmatrix},$$

while the second term enforces W to be row-sparse. Both jointly make X self-represented only by a small number of feature columns in X.

LRSL algorithm [16]: LRSL is an unsupervised feature selection algorithm. This algorithm characterizes the similarity among samples by constructing a graph Laplacian matrix for the addition of structure learning, and it applies $l_{2,\,1}$-norm and rank constraints to the feature coefficient matrix. The optimization problem of the LRSL algorithm is formulated as follows:

$$\begin{aligned} &\min_{W} \frac{1}{2} \|X - XW\|_F^2 + \frac{\alpha}{2} Tr((XW)^T L(XW)) + \beta \|W\|_{2,\,1}, \\ &\text{st.} \quad \operatorname{rank}(W) = k, \end{aligned} \tag{5}$$

where α is the coefficient for adjusting the weight of structural learning, and β is the regularization parameter that controls the row sparsity of W. Additionally, k is a preset constant. The Laplacian matrix L is calculated as $L = D - A$, where the diagonal matrix D is defined as $D_{ii} = \sum_{j=1}^{n} A_{ij}$, and it is normalized as $\mathrm{L} = (I - A)^T (I - A)$. Moreover, the similarity matrix A is calculated as

$$A_{ij} = \begin{cases} e^{-\frac{\|x_i - x_j\|_2^2}{t}}, & x_i \in N_k(x_i),\, x_j \in N_k(x_j) \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where $N_k(x_i)$ is the set of k-nearest neighbors of the sample x_i, and t is the Gaussian kernel parameter which serves to adjust the radial range.

LRSSR Algorithm [17]: This algorithm is an extension of the RSR algorithm by introducing a Schatten p-norm to the feature coefficient matrix W as a low-rank constraint for exploring the internal representational relationship among features. Concurrently, it employs the F-norm to measure the residual term. When $p = 1$, the Schatten p-norm converts to the nuclear norm. The optimization problem for the LRSSR algorithm is formulated as follows:

$$\min_{W} \frac{1}{2} \|X - XW\|_F^2 + \beta \|W\|_{2,\,1} + \lambda \|W\|_*, \tag{7}$$

where β and λ are regularization parameters that are employed to control the row sparsity and low-rank degree of W, respectively.

LRFS Algorithm [18]: This algorithm is a supervised feature selection approach based on low-rank constraints proposed for linear regression models. The optimization problem for the LRFS algorithm is formulated as follows:

$$\min_{W} \|Y - XW\|_F^2 + \gamma \|W\|_F^2 + \lambda \|W\|_*, \tag{8}$$

where γ and λ are regularization parameters that are utilized to control the F-norm and the low-rank degree of W, respectively.

JLRDLP Algorithm [19]: It is a spectral feature selection algorithm that integrates low-rank decomposition and locality preservation. It makes use of spectral clustering to predict pseudo-labels F. Moreover, it ensures the sparsity of the feature selection matrix through $l_{2,\,1}$-norm minimization constraints. The feature coefficient matrix W is decomposed into the product of two low-rank matrices to reveal the global correlations among sample features. Its optimization problem is formulated as follows:

$$\begin{aligned} &\min_{A,\,B,\,F} \frac{1}{2} \|F - XAB\|_F^2 + \alpha Tr\left((XAB)^T L(XAB)\right) + \beta \|AB\|_{2,\,1} + \gamma Tr\left(F^T LF\right), \\ &st. \quad F^T F = I_c,\ F \geq 0, \end{aligned} \tag{9}$$

where β is the regularization parameter which is employed to control the degree of sparsity, and α, γ is the coefficient that is utilized to adjust the weight of structural learning.

The objective function of JLRDLP integrates several signal concepts to achieve robust feature selection:

(1) Data fidelity term $\left(\frac{1}{2} \|F - XAB\|_F^2\right)$: This term ensures that the low-rank representation XAB closely approximates the pseudo-label matrix F. It forces the selected features (through A and B) to be predictive of the sample's cluster assignment.

(2) Manifold smoothness terms $\left(\alpha Tr\left((XAB)^T L(XAB)\right)\right.$ and $\left.\gamma Tr\left(F^T LF\right)\right)$: These two terms are based on the graph Laplacian L, which encodes the local geometric structure of the data. The first term forces that if two original samples x_i and x_j are similar (i.e., connected in the graph), their low-rank representations XAB should also be similar. The second term imposes the same smoothness constraint directly on the pseudo-labels F. Together, they ensure that the local neighborhood structure of the THz spectral data is preserved in both the feature representation and the label space, which is often a valid assumption for smoothly varying spectral signals.

(3) Sparsity regularization ($\beta\|AB\|_{2,1}$): As explained for Eq. (3), this term promotes row-sparsity in the combined matrix AB. Since AB acts as the ultimate feature projection/selection matrix, this term directly selects a sparse set of features by driving the rows corresponding to irrelevant frequencies to zero.

(4) Low-rank decomposition ($W = AB$): Factorizing W into two low-rank matrices A and B explicitly enforces a low-rank constraint on the feature transformation. This captures the global correlation among THz spectral features, implying that the selection of features is not independent but governed by a few underlying factors.

3. LRGSL algorithm

In this section, we first present the motivation underlying the LRGSL algorithm. Then, we expound upon the objective function of the algorithm in detail. Finally, we employ the ADMM to optimize and resolve its objective function.

3.1 *Motivation*

In the field of machine learning and pattern recognition, feature selection serves as a crucial step in data processing, as it not only enhances model performance but also reduces model complexity and improves interpretability. Regularized representation, a widely used feature selection paradigm, aims to optimize model generalization by introducing tailored constraints into the objective function-yet existing methods still face limitations when applied to Terahertz Time-Domain Spectroscopy (THz-TDS) data for PCR storage year identification.

• Unsupervised methods (e.g., LRSL, LRFS) fail to leverage label information of PCR storage years, leading to a disconnect between selected features and the target classification task-this is critical because the subtle spectral differences between different-year PCR are only meaningful when linked to label semantics.

• Fixed-constraint designs (e.g., LRSL's uniform low-rank penalty, LRFS's global graph sparsity) cannot address the "band-specific importance" of THz-TDS data: for example, the 0.6-0.8 THz band (corresponding to active components like hesperidin in PCR) requires finer feature preservation, while the 0.3-0.5 THz band (dominated by baseline noise) needs stronger regularization to avoid interference.

• JLRDLP, despite integrating graph Laplacian constraints, lacks a mechanism to maintain consistency between feature selection results and sample labels, resulting in suboptimal performance for supervised classification tasks like PCR storage year discrimination.

In the field of machine learning and pattern recognition, feature selection serves as a crucial step in data processing, as it not only enhances the model performance but also reduces model complexity and improves their interpretability. Regularized representation, being one of the feature selection approaches, aims to reduce the complexity of models by introducing tailored constraints into the objective function, thereby enhancing their generalization capability.

Inspired by methods such as RSR, LRSL, LRSSR, LRFS, and JLRDLP, these models utilize $l_{2,1}$-norm and nuclear norm constraints (Schatten p-norm when $p = 1$) as sparse and low-rank constraints, which are conducive to excavating the inherent sparse and low-rank structures within the data. This clearly indicates that sparse and low-rank constraints play a significant role in feature selection models. Furthermore, the LRSL algorithm and the JLRDLP algorithm provide us with inspiration that through the introduction of graph Laplacian structural constraints, we can capitalize on the structural information of the inherent similarity of the data to enhance the performance of the model, suggesting that the graph Laplacian structure is also a crucial regularization constraint. While the majority of the aforesaid feature selection methods are aimed at unsupervised classification issues, the identification of the storage years of PCR pertains to a supervised classification problem. In order to fully exploit the label information of samples and integrate the advantages of the aforementioned feature selection models, this paper proposes a supervised feature selection algorithm based on sparse low-rank and graph structural learning, and then combines THz-TDS technology to identify PCR from different years.

3.2 *Objective function*

The objective function of the LRGSL algorithm is as follows:

$$\min_{W} \frac{1}{2} \|Y - XW\|_F^2 + \alpha Tr\left((XW)^T L(XW)\right) + \beta \|W\|_{2,1} + \lambda \|W\|_*, \tag{10}$$

here, $W = [w_1;\ w_2;\ \cdots;\ w_d] \in R^{d \times c}$ is the feature coefficient matrix of X, which reflects the significance of features, $L \in R^{n \times n}$ is the graph Laplacian matrix, consistent with the graph Laplacian in the LRSL algorithm, and it can represent the similarity among samples. The regularization parameters β and λ are employed to control the row sparsity and low-rank of W. The goal is to explore the relationships among feature points by imposing sparse and low-rank constraints, utilizing the $l_{2,1}$-norm and nuclear norm to obtain a sparse coefficient matrix, thereby eliminating redundant features and achieving the purpose of dimensionality reduction. Furthermore, the Laplacian graph structure is employed to maintain the original similarity among samples subsequent to the dimensionality reduction process.

In line with traditional techniques for feature selection based on filtering and wrapper methods, the LRGSL method selects significant features in accordance with the feature coefficient matrix. Nevertheless, its distinctiveness resides in the fact that it takes into comprehensive consideration the sparsity, the low-rank characteristic of the feature coefficient matrix, and the similarity of sample structures throughout the optimization procedure.

3.3 *Optimization*

Here, we employ the ADMM algorithm to address the optimization problem in Eq. (10). Firstly, through the introduction of auxiliary variables A and B, the Eq. (10) is transformed into the ensuing equivalent optimization problem:

$$\begin{aligned} &\min_{W} \frac{1}{2} \|Y - XW\|_F^2 + \alpha Tr\left((XW)^T L(XW)\right) + \beta \|W\|_{2,1} + \lambda \|W\|_* \\ &st. \quad A = W;\ B = W. \end{aligned} \tag{11}$$

Eq. (11) can be solved by constructing the Lagrangian function of the minimization problem, which is expressed in Eq. (12) as follows:

$$\begin{aligned} L(W,\ A,\ B) = &\frac{1}{2} \|Y - XW\|_F^2 + \alpha Tr\left((XW)^T L(XW)\right) + \beta \|A\|_{2,1} \\ &+ \lambda \|B\|_* + \frac{\mu}{2} \left[\|A - W + Y_1\|_F^2 + \|B - W + Y_2\|_F^2\right], \end{aligned} \tag{12}$$

where μ represents the non-negative penalty parameter, and Y_1, Y_2 represents the Lagrange multiplier.

Eq. (12) can be solved via a variable-fixing strategy where variables A, B and W are held constant, then updating the Lagrange multipliers Y_1, Y_2 and μ.

3.3.1 *Fix A and B, update W*

When A and B are fixed, W can therefore be obtained by minimizing the function in Eq. (13),

$$L(W) = \frac{1}{2} \|Y - XW\|_F^2 + \alpha Tr\left((XW)^T L(XW)\right) + \frac{\mu}{2} \left[\|A - W + Y_1\|_F^2 + \|B - W + Y_2\|_F^2\right]. \tag{13}$$

By taking the partial derivative of Eq. (13) w.r.t. W and set it to zero, then we can obtain Eq. (14),

$$W = \left(X^T X + 2\alpha X^T L X + 2\mu I\right)^{-1} \left(X^T Y + \mu(A + B + Y_1 + Y_2)\right), \tag{14}$$

where $I \in R^{d\times d}$ is the identity matrix.

3.3.2 *Fix W and B, update A*

When B is fixed, W can be obtained from Eq. (14). Therefore, A can be determined by minimizing the function in Eq. (15),

$$L(A) = \beta \|A\|_{2,\,1} + \frac{\mu}{2} \|A - W + Y_1\|_F^2. \tag{15}$$

According to the soft-thresholding shrinkage theorem, A can be represented by a row-wise shrinkage function of matrix $D = W - Y_1$ with respect to the threshold $\dfrac{\beta}{\mu}$. Therefore, the iterative formula A is given by Eq. (16),

$$A = \text{mat-soft}\left(D,\, \frac{\beta}{\mu}\right) = \begin{pmatrix} \text{vect-soft}\left(D_1,\, \dfrac{\beta}{\mu}\right) \\ \vdots \\ \text{vect-soft}\left(D_d,\, \dfrac{\beta}{\mu}\right) \end{pmatrix}, \tag{16}$$

where

$$\text{vect-soft}\left(D_i,\, \frac{\beta}{\mu}\right) = \frac{\max\left\{\|D_i\|_2 - \dfrac{\beta}{\mu},\, 0\right\}}{\max\left\{\|D_i\|_2 - \dfrac{\beta}{\mu},\, 0\right\} + \dfrac{\beta}{\mu}} D_i,$$

D_i $(i = 1,\, \cdots,\, d)$ be the i-th row of D.

3.3.3 *Fix W and A, update B*

As W and A can be obtained from the iterative Eq. (14) and Eq. (16). Then, B can be derived through minimizing the function in Eq. (17),

$$L(B) = \lambda \|B\|_* + \frac{\mu}{2} \|B - W + Y_2\|_F^2. \tag{17}$$

Based on the soft-thresholding shrinkage theorem, we can derive the iterative formula (18) for B,

$$B = U S_{\left(\frac{\lambda}{\mu}\right)}(\Sigma) V^T. \tag{18}$$

Where $W - Y_2 = U\Sigma V^T$, $S_{\left(\frac{\lambda}{\mu}\right)}(\Sigma) = \text{diag}\left(\max\left(\Sigma^{ii} - \frac{\lambda}{\mu}, 0\right)\right)$, and Σ^{ii} be the i-th singular value of diagonal matrix Σ.

3.3.4 *Update Y_1, Y_2*

W, A, and B are iteratively updated using Eq. (14), (16), and (18), respectively. The iterative update formulas for the Lagrange multipliers Y_1 and Y_2 are given by Eq. (19) and (20) as follows:

$$Y_1 = Y_1 + A - W, \tag{19}$$

$$Y_2 = Y_2 + B - W, \tag{20}$$

$$\mu = \rho\mu, \tag{21}$$

where $\rho \in R$ is a constant, which is set to 1.5 in the experiments of this paper.

Based on the previous discussion and the resolution of all optimization variables related to the LRGSL feature selection issue, the feature selection process of the LRGSL algorithm is presented in Algorithm 1.

Algorithm 1 LRGSL Algorithm

Input: Data matrix $X \in R^{n\times d}$; Label matrix $Y \in R^{n\times c}$; Laplacian matrix $L \in R^{n\times n}$; number of selected features $f \leq d$; parameters α, β, λ, maximum number of iterations K, error tolerance ε.

1. Initialization: $W_0 = A_0 = B_0 = Y_{1,\,0} = Y_{2,\,0} = 0$, $k = 0$, $\mu_0 = \lambda/0.01$, $\rho = 1.5$, $k = 0$, $err = \inf$.
2. **while** $err > \varepsilon$ and $k < K$ **do**
3. Update W by solving Eq. (14);
4. Update A via row-wise shrinkage (Eq. (16));
5. Update B via singular value thresholding (Eq. (18));
6. Update Y_1, Y_2, μ (Eq. (19)-(21));
7. Compute error $err = \dfrac{\|W_{k+1} - W_k\|_F}{\|W_{k+1}\|_F} < \varepsilon$;
8. $k = k + 1$;
9. **end while**
10. Rank features by l_2-norm of rows of W and select top f features.

Output: Feature subset Ω_{opt}.

4. Experimental preparation and methodology

4.1 *Datasets*

The THz spectral dataset of PCR used in this study is consistent with that employed in our previous work [20]. Detailed information regarding the THz-TDS instrument and the collection of PCR's THz spectral data can be found in the relevant descriptions of [20].

The dataset consists of four categories classified by the storage year of PCR, specifically 1-year, 3-year, 5-year, and 7-year. A detailed overview of the dataset is provided in Table 1. Six spectral measurements were acquired from each independent batch of raw materials; consequently, the total sample size for each category is the product of the number of batches and six. The dataset comprises a total of 162 batches, yielding 972 spectral samples. The distribution across the four age categories is as follows: 54 batches (324 samples) for 1-year, 43 batches (258 samples) for 3-year, 36 batches (216 samples) for 5-year, and 29 batches (174 samples) for 7-year. This distribution reflects the practical challenge of acquiring older samples, which are naturally less common and more valuable. The imbalanced nature of the dataset is a common

scenario in real-world classification tasks, particularly with aged biological specimens. In our subsequent analysis and modeling, we took this imbalance into consideration to ensure that the developed feature selection and classification methods are robust and not biased towards the majority classes (e.g., 1-year and 2-year).

Table 1. PCR dataset overview-category labels, batch counts, and sample sizes

Data	PCR			
Label	1	3	5	7
Batches	54	43	36	29
Samples	324	258	216	174

Figure 1 presents the THz spectra of PCR in the 0.1-2 THz frequency range. The Figure 1a shows the THz absorption coefficient spectra of all samples, while the Figure 1c displays the corresponding refractive index spectra. The Figure 1b illustrates the average absorption coefficient spectra of PCR samples stored for 1, 3, 5, and 7 years, and the Figure 1d correspondingly shows their average refractive index spectra. In each subplot, the THz data of PCR samples from the same storage year are represented by lines of the same color.

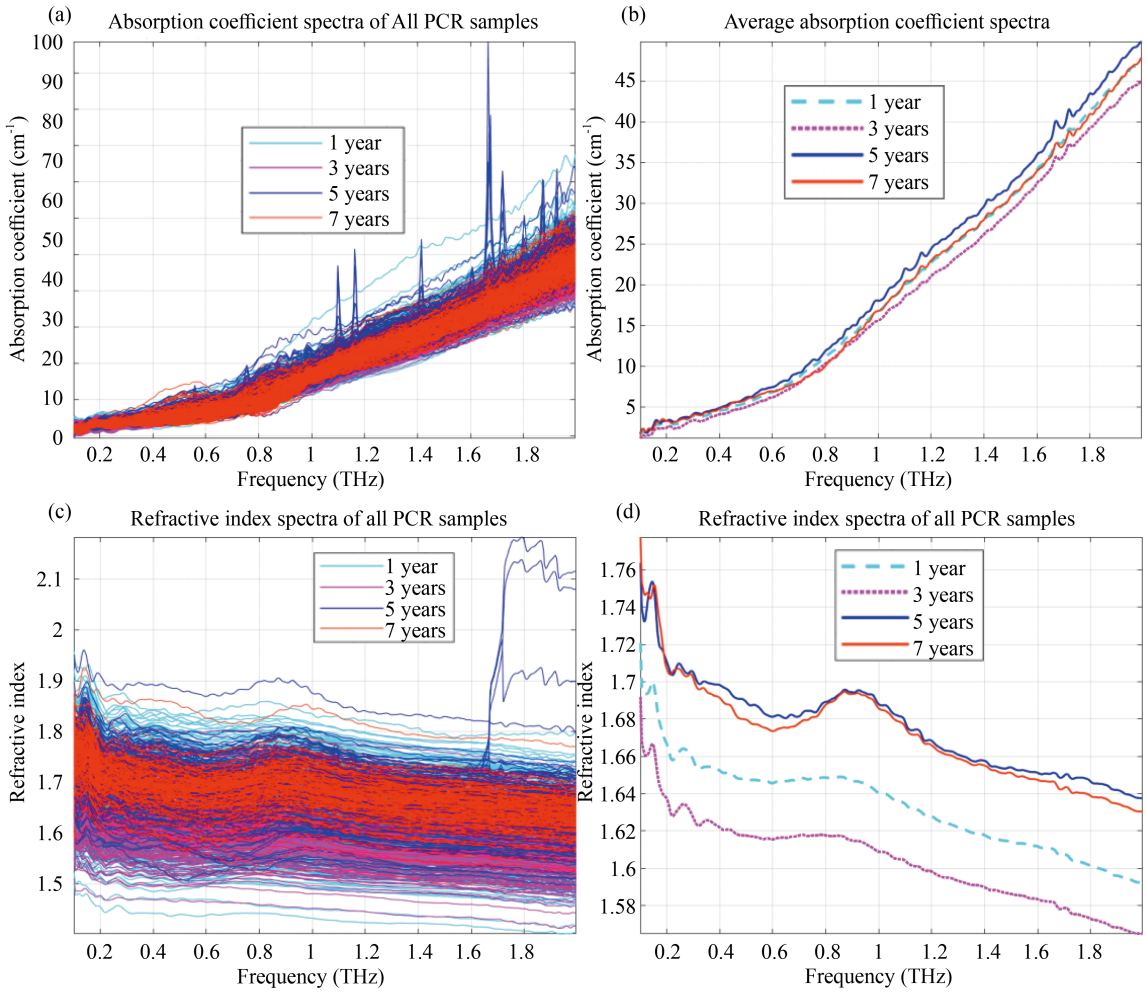


Figure 1. Terahertz absorption coefficient and refractive index spectra of PCR stored for different years in the 0.1-2 THz range

As can be seen from Figure 1a, c, the spectra of different PCR samples are mixed together without showing clear stratification. The absorption coefficient spectra exhibit obvious anomalies beyond 1 THz, and some of the refractive index spectra also have anomalies beyond 1.5 THz. From the average spectral curves of each category in Figure 1b, d, the spectral data of different PCR samples show a certain degree of stratification. For example, the average absorption coefficient and refractive index of 3-year PCR samples are the smallest, while those of 5-year samples are relatively higher. The average absorption coefficient spectra of 1-year and 7-year PCR samples are difficult to distinguish, and the average refractive index spectra of 5-year and 7-year samples are quite close. Figure 1 indicates that it is impossible to directly distinguish PCR samples of different storage years solely by observing the spectral curves, and machine learning methods combined with feature selection have emerged as a solution for this purpose.

4.2 *Experimental setting*

Although PCR samples from different storage years exhibit compositional complexity, their THz spectra show highly similar curve shapes. In contrast, the differences in absorption coefficient and refractive index spectra among PCR samples of different storage years are more pronounced than those observed in time-domain and frequency-domain spectra [21]. Zhan et al. [22] successfully combined absorption coefficient and refractive index spectra with chemometric methods for the rapid identification of crude oils from different geographical origins, achieving favorable prediction performance. Therefore, within appropriate frequency ranges, absorption coefficient and refractive index spectra can be effectively used for qualitative discrimination of PCR samples with different aging years. Given the anomalies in the absorption coefficient spectra beyond 1 THz and the refractive index spectra beyond 1.5 THz (see Figure 1), this study will select the absorption coefficient spectra in the range of 0.1-1 THz and the refractive index spectra in the range of 0.1-1.5 THz for testing.

Given the high similarity among multiple pellet samples compressed from the same batch, adopting a conventional random sample partitioning approach may lead to data leakage issues [23, 24]. To mitigate the influence of experimental randomness and align the experimental design more closely with real-world engineering applications, a batch-based sample partitioning strategy was employed to divide the dataset into training and test sets. Specifically, 25 batches from each class were randomly selected to form the training set, and 4 additional batches were randomly chosen from the remaining batches as the test set. To avoid the contingency of a single experiment, the batch-based partitioning was repeated 10 times. For a given algorithm, the parameter combination that yielded the highest average recognition accuracy over these 10 experiments was considered optimal. We report the average Accuracy (ACC) and Standard Deviation (STD) of each algorithm under its optimal parameters across the 10 experiments.

The optimal parameters for all algorithms were determined via grid search. The regularization parameters α, β, λ and γ were searched over the set $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$; the low-rank parameter k (used in the LRSL algorithm) was searched within the set $\{1, 3, 5, 7, 9\}$. The maximum number of iterations K and error tolerance ε are set to 200 and 10^{-4} in all experiments of this paper.

We utilized a SVM to assess the discriminative power of the selected features identified by each regularized representation-based method, with the baseline defined as the SVM classification accuracy on the full spectrum without feature selection. For the given spectral dataset, without feature selection, grid search was first used to obtain the SVM parameters that achieved the highest average accuracy in 10 Monte Carlo experiments. This accuracy served as the baseline, and subsequent feature selection algorithms all adopted these SVM parameters for training and testing. Specifically, no separate SVM parameter optimization was performed during the testing of feature selection algorithms; otherwise, the combined parameters of the feature selection algorithms and SVM would result in an excessively large number of parameters to tune. It is noted here that for both the absorption coefficient spectral data and the refractive index spectral data, consistent results were obtained in the SVM parameter search: the kernel function was the linear kernel for all cases, the data were standardized uniformly, the Box Constraint parameter was mostly in the range of 0.1-10, and only a few were in the range of 20-30.

5. Experimental results and analysis

5.1 *Performance comparison of feature selection methods*

As the foundation of subsequent analysis, this subsection first focuses on the core task of PCR year classification and compares the overall performance of six feature selection methods (RSR, LRSSR, LRFS, LRSL, JLRDLP, and LRGSL) on two standard spectral datasets: absorption coefficient spectra (0.1-1.0 THz) and refractive index spectra (0.1-1.5 THz). The classification ACC with corresponding standard deviations is used as the evaluation metric.

Table 2 shows the best ACC and STD of each method, where Baseline refers to the straightforward application of SVM for classification, bypassing any feature selection process.

Table 2. The performance of different feature selection methods on absorption coefficient spectra (0.1-1.0 THz) and refractive index spectra (0.1-1.5 THz): ACC ± STD (%)

Methods	Absorption coefficient spectra	Refractive index spectra
Baseline	75.29 ± 5.71	74.46 ± 5.51
RSR	75.39 ± 4.86	74.56 ± 5.76
LRSSR	75.91 ± 5.18	75.20 ± 5.61
LRFS	75.91 ± 5.21	75.19 ± 6.21
LRSL	75.39 ± 5.84	74.77 ± 5.99
JLRDLP	75.91 ± 4.99	74.77 ± 6.17
LRGSL	**76.22 ± 5.23**	**75.29 ± 5.82**

Table 2 presents the classification performance (expressed as ACC ± STD in percentage) of various feature selection methods under two spectral data scenarios: absorption coefficient spectra and refractive index spectra. Among all the compared methods, the proposed LRGSL method achieves the best performance in both spectral datasets, with ACC values of 76.22 ± 5.23% for absorption coefficient spectra and 75.29 ± 5.82% for refractive index spectra. The LRSSR, LRFS, and JLRDLP methods exhibit comparable performance, yielding ACC values around 75.91% for absorption coefficient spectra and approximately 74.77-75.20% for refractive index spectra. In contrast, the Baseline and RSR methods show relatively lower classification accuracy, with their ACC values remaining below 75.40% across both spectral datasets. These results demonstrate that the LRGSL method is more effective in capturing discriminative features from spectral data compared to other state-of-the-art feature selection methods, and further confirm the superior discriminative capability of absorption coefficient spectra.

5.2 *Impact of feature number on classification accuracy*

This subsection presents the variation of the optimal classification accuracy of the six feature selection algorithms under different numbers of selected features.

Figure 2 depicts the ACCs of the feature selection methods in identifying the storage year of PCR across various feature dimensions. As shown in Figure 2, overall, the ACC of the feature selection methods progressively increases along with the augmentation of feature dimensions. When the number of features for the absorption coefficient spectrum and the refractive index spectrum is both set to 150, the LRGSL method achieves the highest accuracies (76.22% and 75.29%), respectively. More specifically, firstly, compared to the RSR method with sparse constraints, the proposed LRGSL method exhibits superior performance across diverse feature dimensions. This is because the low-rank constraint can effectively exploit the inherent low-rank properties of the data and the correlations among features, thereby enhancing the accuracy and efficiency of the feature selection method. Secondly, compared to the LRSSR method with sparse and low-rank constraints, LRGSL utilizes the Laplacian graph structure to maintain the original similarity among samples after dimensionality reduction, thereby resulting in a further improvement in the ACC of the LRGSL method. This verifies that the combination of low-rank and graph structure constraints enables the feature selection model to further capture feature information, thereby enhancing the performance of the model and contributing to a reduction in computational space.

Therefore, the combination of THz-TDS technology and feature selection methods based on regularized representations is demonstrably feasible for the identification of the storage years of PCR.

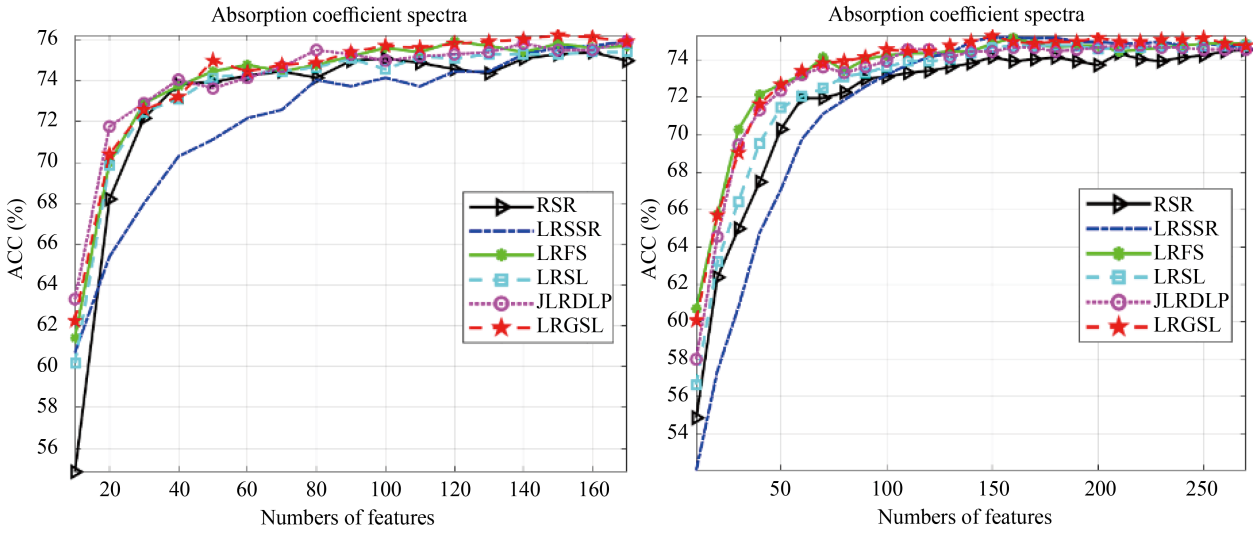


Figure 2. The ACC of feature selection methods under varying numbers of features

5.3 *Analysis of top 50 important features*

This subsection presents the top 50 features selected by each of the six algorithms on both the absorption coefficient spectra and refractive index spectra when they achieve the highest classification accuracy. The number of 50 is chosen instead of a larger value because a excessively large number would result in the selected features including many less important ones, making it difficult to truly identify the specific frequency bands where the important features chosen by each algorithm are concentrated.

Figure 3 presents these results. The *x*-axis represents frequency (in THz), while the y-axis indicates different feature selection algorithms. For the absorption coefficient spectra, the features selected by the RSR algorithm are mainly concentrated in the ranges of 0.1-0.15 THz, 0.3-0.6 THz, and 0.7-0.8 THz. In contrast, the features chosen by the LRSSR algorithm are primarily distributed within 0.1-0.23 THz and 0.85-1 THz. The feature distributions selected by the LRFS, JLRDLP, and LRGSL algorithms exhibit a certain degree of similarity, characterized by discrete, short segments. Meanwhile, the features selected by the LRSL algorithm show an overall uniform distribution. In general, the 0.1-0.2 THz range is consistently selected by all feature selection algorithms, indicating that this frequency band contains important discriminative features for classification. For the refractive index spectra, the distribution of features selected by the RSR algorithm is relatively similar to that in the absorption coefficient spectra, both being scattered. The features chosen by the LRSSR algorithm are only concentrated in the 0.1-0.3 THz range. Similarly, the LRFS algorithm primarily selects features within 0.1-0.3 THz but also includes a small number of points in the 0.6-1 THz range. The LRSL algorithm selects features mainly in the ranges of 0.1-0.2 THz, around 0.4 THz, 0.6-0.7 THz, and 0.9-1 THz. The JLRDLP algorithm focuses on features near 0.1 THz, 0.4-0.7 THz, around 0.9 THz, and around 1.5 THz. For the LRGSL algorithm, the selected features are mainly concentrated in 0.1-0.2 THz, around 0.7 THz, and 0.9-1 THz. Overall, the 0.1-0.2 THz range is also the most frequently selected band by all feature selection algorithms, confirming its significance for classification. In summary, whether for absorption coefficient spectra or refractive index spectra, the discriminative information related to the storage years of PCR is likely concentrated primarily in the lowest frequency region.

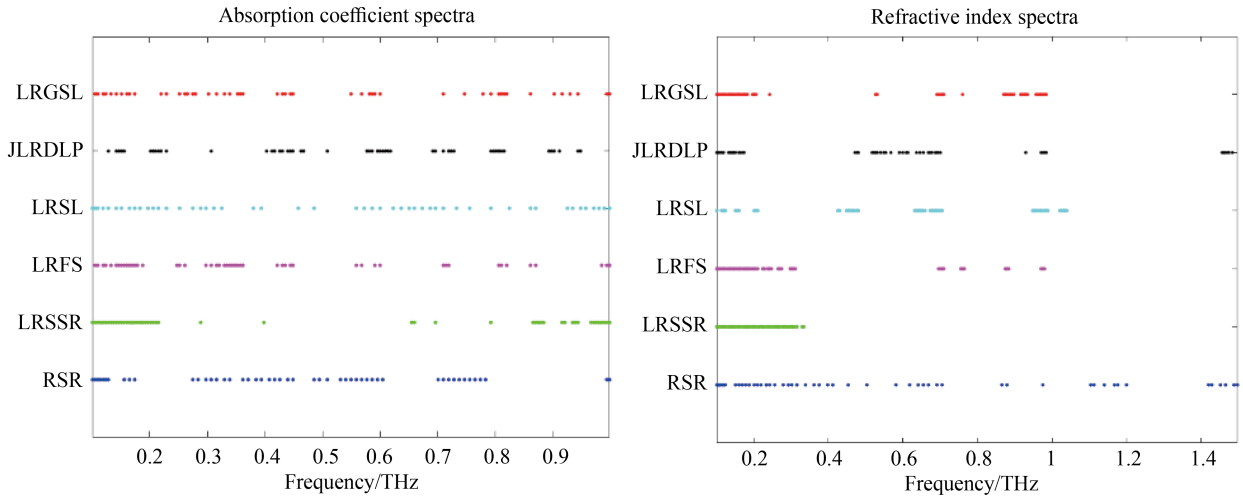


Figure 3. Top 50 selected features of six algorithms on absorption coefficient spectra (Left) and refractive index spectra (Right)

5.4 *Supplementary experiment: performance under adjusted frequency ranges*

To verify the robustness of the proposed LRGSL method and existing feature selection algorithms against variations in spectral frequency ranges, a supplementary experiment was conducted. Unlike the standard frequency ranges in Table 2 (absorption coefficient spectra: 0.1-1.0 THz; refractive index spectra: 0.1-1.5 THz), the frequency ranges here were adjusted to absorption coefficient spectra (0.1-1.5 THz) and refractive index spectra (0.1-1.0 THz), with classification accuracy (ACC ± STD, %) as the evaluation metric. The results are summarized in Table 3.

A comparison between Table 2 and Table 3 reveals two key observations:

1) Overall performance trend consistency: Despite the adjusted frequency ranges, the relative performance ranking of the six feature selection algorithms remains largely consistent with the standard setting. Specifically, LRGSL still achieves the best overall performance-securing the highest accuracy (75.08 ± 5.08%) on the extended absorption coefficient spectrum (0.1-1.5 THz) and sharing the top spot (74.25 ± 5.75%) on the truncated refractive index spectrum (0.1-1.0 THz) with LRSSR. This confirms that the superiority of LRGSL is not dependent on specific frequency range configurations, reflecting good robustness.

2) Sensitivity to frequency range extension/truncation: For most methods (including the Baseline), the classification accuracy on the absorption coefficient spectrum decreases slightly when the frequency range is extended from 0.1-1.0 THz to 0.1-1.5 THz (e.g., Baseline drops from 75.29 ± 5.71% to 72.26 ± 4.97%). This may be attributed to the introduction of redundant or noisy spectral information in the high-frequency region (1.0-1.5 THz) that is irrelevant to PCR storage year classification. Conversely, truncating the refractive index spectrum from 0.1-1.5 THz to 0.1-1.0 THz leads to a mild performance decline for most algorithms, indicating that the high-frequency segment (1.0-1.5 THz) of the refractive index spectrum may contain weak but non-negligible discriminative information.

Notably, LRGSL exhibits the smallest performance degradation among all methods when the frequency ranges are adjusted, further demonstrating its ability to adapt to variations in spectral data distribution and filter out irrelevant information effectively. Collectively, the supplementary experiment validates that the proposed LRGSL method and the comparative algorithms maintain stable performance trends under adjusted frequency ranges, and LRGSL's robustness makes it more suitable for practical PCR quality evaluation scenarios where spectral frequency ranges may vary slightly.

Table 3. Classification accuracy of different methods on absorption coefficient spectra (0.1-1.5 THz) and refractive index spectra (0.1-1.0 THz): ACC ± STD (%)

Methods	Absorption coefficient spectra	Refractive index spectra
Baseline	72.26 ± 4.97	73.83 ± 5.79
RSR	74.35 ± 5.60	73.83 ± 5.58
LRSSR	74.45 ± 4.66	**74.25 ± 5.73**
LRFS	74.87 ± 5.46	74.15 ± 5.68
LRSL	74.56 ± 5.45	74.04 ± 5.77
JLRDLP	74.87 ± 4.29	74.04 ± 5.65
LRGSL	**75.08 ± 5.08**	**74.25 ± 5.75**

5.5 *Parameter sensitivity analysis of LRGSL algorithm*

In Table 2, the LRGSL algorithm achieves the highest accuracy of 76.22 ± 5.23% and 75.29 ± 5.82% on the absorption coefficient spectra and refractive index spectra, respectively. The corresponding optimal parameters are $\alpha = 0.1$, $\lambda = 10^{-3}$, $\beta = 0.01$ and $\alpha = 1$, $\lambda = 10^{-4}$, $\beta = 1$, with 150 features used in both cases. To analyze the parameter sensitivity of the LRGSL algorithm, we fix the number of features at 150 and then present the variations in the average classification accuracy of LRGSL across 10 Monte Carlo tests when one of the parameters (α, λ, or β) is fixed while the other two are varied.

Figures 4a-c show the test results on the absorption coefficient spectra dataset when α, λ, and β are fixed, respectively. Figures 4d-f display the corresponding results on the refractive index spectra dataset when α, λ, and β are fixed, respectively. As observed from the figures, the changes in ACC are relatively small in all cases, indicating that the LRGSL algorithm is insensitive to parameter variations when selecting effective features.

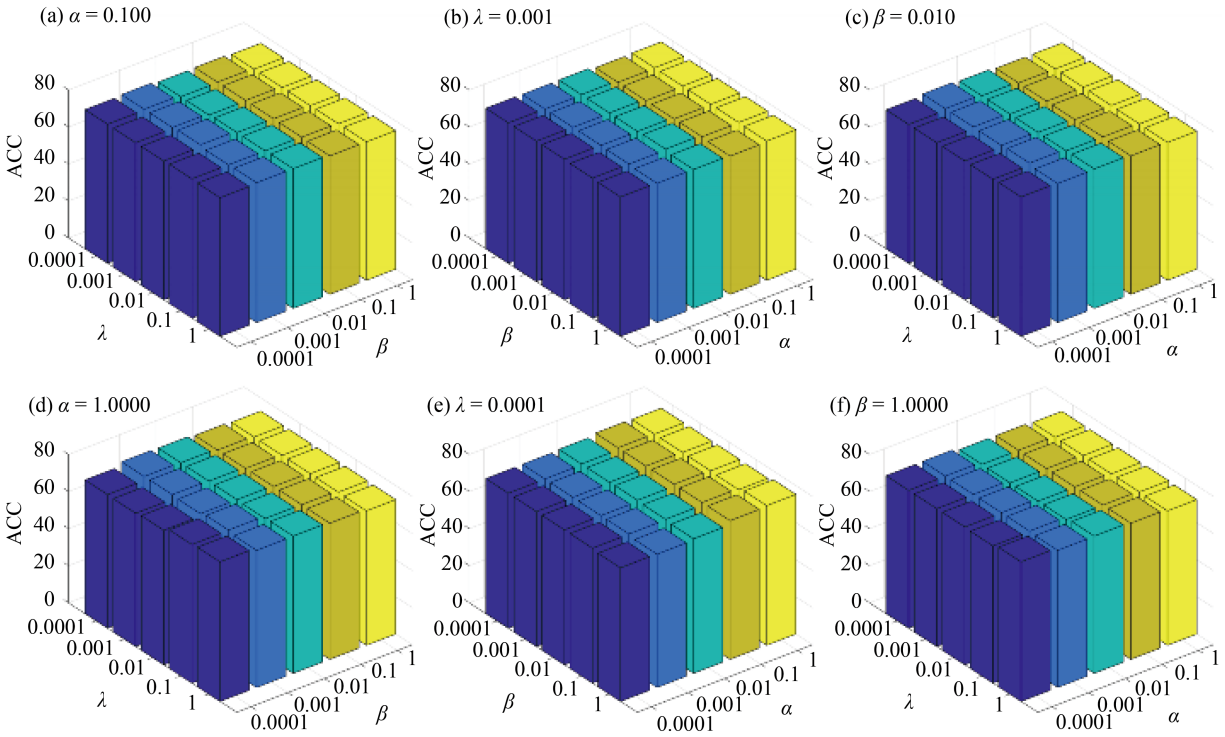


Figure 4. Parameter sensitivity snalysis of LRGSL: ACC variations with α, λ, and β (top row: sbsorption coefficient spectra; bottom row: refractive index spectra)

6. Conclusion

This paper has investigated the feasibility of integrating THz-TDS with regularized representation-based feature selection methods for identifying the storage years of PCR. Several established feature selection techniques, including RSR, LRSSR, LRFS, LRSL, and JLRDLP, were applied and objectively evaluated for this task. Building upon their strengths, a novel feature selection method based on LRGSL was proposed, which explores the inherent low-rank data structure via sparse and low-rank constraints on the feature matrix while incorporating sample similarities through a Laplacian graph.

Experimental results confirm that the combination of THz-TDS technology and regularized feature selection methods provides an effective for identifying PCR storage years. The incorporation of sparse and low-rank constraints was shown to enhance feature selection performance. It was also observed that the efficacy of these methods is highly dependent on the THz frequency range utilized. Among the methods evaluated, LRGSL, JLRDLP, and LRSL demonstrated superior performance.

The practical implication of this work lies in its potential application for quality control of Traditional Chinese Medicines (TCMs), such as PCR, where non-destructive, rapid, and accurate determination of storage duration is critical for ensuring medicinal quality and consistency.

However, this study has certain limitations. The evaluation was conducted on a single, application-specific dataset. While the batch-wise validation strategy was rigorous, the overall sample size remains moderate, and the absolute gains in accuracy are modest. Future work will focus on validating the generalizability of this approach on larger datasets and other TCMs. Furthermore, we will explore advanced deep learning-based feature selection methods and seek to establish a more explicit physical interpretation between the selected THz spectral features and the biochemical changes in aging PCR.

Acknowledgement

This work was supported by the Wuyi University Hong Kong-Macau joint Research and Development Fund under Grant 2022WGALH16, supported in part by the Science and Technology Development Fund of Macau SAR under Grant 0079/2023/ITP2 and 0097/2025/ITP2, the Special Foundation in Key Fields for Universities of Guangdong Province under Grant 2023ZDZX4060 and the Guangdong Province Education Science Planning Project (Higher Education Special) under Grant 2024GXJK275.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Fu M, Zou B, An K, Yu Y, Tang D, Wu J, et al. Anti-asthmatic activity of alkaloid compounds from Pericarpium Citri Reticulatae (*Citrus reticulata* 'Chachi'). *Food and Function*. 2019; 10(2): 903-911. Available from: https://doi.org/10.1039/C8FO01753K.

[2] Yu X, Sun S, Guo Y, Liu Y, Yang D, Li G, et al. Citri Reticulatae Pericarpium (Chenpi): botany, ethnopharmacology, phytochemistry, and pharmacology of a frequently used traditional Chinese medicine. *Journal of Ethnopharmacology*. 2018; 220: 265-282. Available from: https://doi.org/10.1016/j.jep.2018.03.031.

[3] Chen L, Xiao G, Xu Y, Cheang WS, Wu J, Yu Y, et al. Effects of aging on the bioactive flavonoids and fungal diversity of Pericarpium Citri Reticulatae (*Citrus reticulata* 'Chachi'). *LWT*. 2024; 200: 116135. Available from: https://doi.org/10.1016/j.lwt.2024.116135.

[4] Husni E, Hamidi D, Pavvellin D, Hidayah H, Syafri S. Metabolite profiling, antioxidant, and in vitro wound healing activities of Citrus medica L. and Citrus *x* microcarpa Bunge peels and leaves essential oils. *Prospects in Pharmaceutical Sciences*. 2024; 22(4): 122-130. Available from: https://doi.org/10.56782/pps.250.

[5] Li F, Lu Y, Li C, Huang R, Tian E, Tan E, et al. *trnL-trn*F copy number is inversely correlated with storage time of Guang Chenpi, the aged sun-dried peels of *Citrus reticulata* 'Chachi'. *Journal of Stored Products Research*. 2022; 97: 101982. Available from: https://doi.org/10.1016/j.jspr.2022.101982.

[6] Luo Y, Zeng W, Huang KE, Li DX, Chen W, Yu XQ, et al. Discrimination of Citrus reticulata Blanco and *Citrus reticulata 'Chachi'* as well as the *Citrus reticulata 'Chachi'* within different storage years using ultra high performance liquid chromatography quadrupole/time-of-flight mass spectrometry based metabolomics approach. *Journal of Pharmaceutical and Biomedical Analysis*. 2019; 171: 218-231. Available from: https://doi.org/10.1016/j.jpba.2019.03.056.

[7] Wang H, Chen G, Fu X, Liu RH. Effects of aging on the phytochemical profile and antioxidative activity of *Pericarpium Citri Reticulatae* 'Chachiensis'. *RSC Advances*. 2016; 6(107): 105272-105281. Available from: https://doi.org/10.1039/C6RA22082G.

[8] He X, Huang S, Wu M, Wu M, Zhang Y, Ma Z, et al. Simultaneous quantitative analysis of ten bioactive flavonoids in Citri Reticulatae Pericarpium Viride (Qing Pi) by ultrahigh-performance liquid chromatography and high-resolution mass spectrometry combined with chemometric methods. *Phytochemical Analysis*. 2021; 32(6): 1152-1161. Available from: https://doi.org/10.1002/pca.3056.

[9] Li K, Chen X, Zhang R, Pickwell-MacPherson E. Classification for glucose and lactose terahertz spectrums based on SVM and DNN methods. *IEEE Transactions on Terahertz Science and Technology*. 2020; 10(6): 617-623. Available from: https://doi.org/10.1109/TTHZ.2020.3013819.

[10] Feng C-H, Otani C. Terahertz spectroscopy technology as an innovative technique for food: current state-of-the-art research advances. *Critical Reviews in Food Science and Nutrition*. 2021; 61(15): 2523-2543. Available from: https://doi.org/10.1080/10408398.2020.1779649.

[11] Shen Y, Li B, Li G, Lang C, Wang H, Zhu J, et al. Rapid identification of producing area of wheat using terahertz spectroscopy combined with chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2022; 269: 120694. Available from: https://doi.org/10.1016/j.saa.2021.120694.

[12] Wang K, Sun DW, Pu H. Emerging non-destructive terahertz spectroscopic imaging technique: principle and applications in the agri-food industry. *Trends in Food Science and Technology*. 2017; 67: 93-105. Available from: https://doi.org/10.1016/j.tifs.2017.06.001.

[13] Miao J, Ping Y, Chen Z, Jin XB, Li P, Niu L. Unsupervised feature selection by non-convex regularized self-representation. *Expert Systems with Applications*. 2021; 173: 114643. Available from: https://doi.org/10.1016/j.eswa.2021.114643.

[14] Yang Y, Shen HT, Ma Z, Huang Z, Zhou X. $l_{2,1}$-norm regularized discriminative feature selection for unsupervised learning. In: *IJCAI International Joint Conference on Artificial Intelligence*; 2011; 2: 1589-1594. Available from: https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-267.

[15] Zhu P, Zuo W, Zhang L, Hu Q, Shiu SCK. Unsupervised feature selection by regularized self-representation. *Pattern Recognition*. 2015; 48(2): 438-446. Available from: https://doi.org/10.1016/j.patcog.2014.08.006.

[16] Wang S, Wang H. Unsupervised feature selection via low-rank approximation and structure learning. *Knowledge-Based Systems*. 2017; 124: 70-79. Available from: https://doi.org/10.1016/j.knosys.2017.03.002.

[17] Wang W, Ma J, Xu C, Zhang Y, Ding Y, Yu S, et al. Low-rank sparse feature selection for image classification. *Expert Systems with Applications*. 2022; 189: 115685. Available from: https://doi.org/10.1016/j.eswa.2021.115685.

[18] Lim H. Low-rank learning for feature selection in multi-label classification. *Pattern Recognition Letters*. 2023; 172: 106-112. Available from: https://doi.org/10.1016/j.patrec.2023.05.036.

[19] Jiang L, Chen X. Spectral feature selection via low rank decomposition and local preservation. In: *2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*. Guangzhou, China: IEEE; 2023. p.518-522.

[20] Zheng C, Zheng H, Ye Z, Li Q. Chenpi years identification based on terahertz spectral data and time series deep learning. *Applied Food Research*. 2025; 5(2): 101463. Available from: https://doi.org/10.1016/j.afres.2025.101463.

[21] Liu Y, Pu H, Li Q, Sun D-W. Discrimination of Pericarpium Citri Reticulatae in different years using terahertz time-domain spectroscopy combined with convolutional neural network. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2023; 286: 122035. Available from: https://doi.org/10.1016/j.saa.2022.122035.

[22] Zhan H, Wu S, Bao R, Ge L, Zhao K. Qualitative identification of crude oils from different oil fields using terahertz time-domain spectroscopy. *Fuel*. 2015; 143: 189-193. Available from: https://doi.org/10.1016/j.fuel.2014.11.047.

[23] Zheng C, Cai S, Li Q, Li C, Li X. A collaborative classification algorithm with multi-view terahertz spectra. *Results in Physics*. 2022; 42: 106023. Available from: https://doi.org/10.1016/j.rinp.2022.106023.

[24] Zheng C, Zha X, Cai S, Cui J, Li Q, Ye Z. Interval-based sparse ensemble multi-class classification algorithm for terahertz data. *Heliyon*. 2024; 10(6): e27743. Available from: https://doi.org/10.1016/j.heliyon.2024.e27743.